

Chalmers Publication Library



Copyright Notice

©2011 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

This document was downloaded from Chalmers Publication Library (<http://publications.lib.chalmers.se/>), where it is available in accordance with the IEEE PSPB Operations Manual, amended 19 Nov. 2010, Sec. 8.1.9 (<http://www.ieee.org/documents/opsmanual.pdf>)

(Article begins on next page)

Noncoherent SIMO Pre-Log via Resolution of Singularities

Erwin Riegler¹, Veniamin I. Morgenshtern², Giuseppe Durisi³,
Shaowei Lin⁴, Bernd Sturmfels⁴, Helmut Bölcskei²

¹Vienna University of Technology, 1040 Vienna, Austria

²ETH Zurich, 8092 Zurich, Switzerland

³Chalmers University of Technology, 41296 Gothenburg, Sweden

⁴University of California, Berkeley, CA 94720

Abstract—We establish a lower bound on the noncoherent capacity pre-log of a temporally correlated Rayleigh block-fading single-input multiple-output (SIMO) channel. Our result holds for arbitrary rank Q of the channel correlation matrix, arbitrary block-length $L > Q$, and arbitrary number of receive antennas R , and includes the result in Morgenshtern et al. (2010) as a special case. It is well known that the capacity pre-log for this channel in the single-input single-output (SISO) case is given by $1 - Q/L$, where Q/L is the penalty incurred by channel uncertainty. Our result reveals that this penalty can be reduced to $1/L$ by adding only one receive antenna, provided that $L \geq 2Q - 1$ and the channel correlation matrix satisfies mild technical conditions. The main technical tool used to prove our result is Hironaka’s celebrated theorem on resolution of singularities in algebraic geometry.

I. INTRODUCTION

It was shown in [1] that the *noncoherent* capacity¹ pre-log for single-input multiple-output (SIMO) correlated block-fading channels can be larger than the pre-log in the single-input single-output (SISO) case. This result was surprising as it disproved a conjecture in [2] on the pre-log in the SIMO case being the same as in the SISO case. The channel model analyzed in [1] assumes that the fading process is independent across blocks of length L and temporally correlated within blocks, with the rank of the corresponding $L \times L$ channel correlation matrix given by $Q < L$ and the number of receive antennas $R = Q$. For this channel model, under additional technical conditions on the channel correlation matrix, a pre-log of $(1 - 1/L)$ was established in [1]. In contrast, in the SISO case the pre-log is given by $1 - Q/L$.

The assumption $R = Q$ made in [1] is very restrictive, and the proof technique used in [1] heavily relies on this assumption. More precisely, the main result in [1] is based on a lower bound on the differential entropy $h(\mathbf{y})$ of the channel output signal that is obtained by applying a change of variables argument [1, Lem. 3]. The proof is then completed

by showing that the expected logarithm of the Jacobian determinant corresponding to this change of variables is finite. For $R < Q$, the Jacobian determinant takes a very involved form making it difficult to say anything about its expected logarithm. The main contribution of this paper is to resolve this problem by introducing a new proof technique based on a result from algebraic geometry, namely [3, Th. 2.3], which is a consequence of Hironaka’s celebrated theorem on resolution of singularities [4], [5]. Roughly speaking, this result allows us to rewrite any real analytic function [6, Def. 1.6.1] locally as a product of a monomial and a *nonvanishing* real analytic function. The proof of our main result, a lower bound on the pre-log for the correlated block-fading channel with *arbitrary* number of receive antennas $R \leq Q$, is then effected by using this factorization to show that the integral of the logarithm of the absolute value of a real analytic function over a compact set is finite, provided that the real analytic function is not identically zero. This method is very general and could be of independent interest when one tries to show that a certain differential entropy is finite.

We conclude by noting that the main result in this paper shows that the pre-log penalty Q/L incurred in the SISO case, which is due to channel uncertainty, can be reduced to $1/L$ by adding only one receive antenna (i.e., by taking $R = 2$), provided that $L \geq 2Q - 1$ and the channel correlation matrix satisfies mild technical conditions. In the limit $L, Q \rightarrow \infty$ with L/Q constant, the $1/L$ penalty in the SIMO case becomes arbitrarily small, whereas the Q/L penalty in the SISO case remains unchanged.

A. Notation

Finite subsets of the set of natural numbers, $\mathcal{I} \subset \mathbb{N}$, are denoted by calligraphic letters and we write $|\mathcal{I}|$ for the cardinality of \mathcal{I} . We use $[m:n]$ to designate the set of natural numbers $\{m, m+1, \dots, n\}$. Uppercase boldface letters denote matrices, lowercase boldface letters designate vectors. The superscripts ^T and ^H stand for transposition and Hermitian transposition, respectively. The all-zero matrix of appropriate size is written as $\mathbf{0}$. For a matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$, the entry in the

The work of Erwin Riegler was partially supported by the WWTF project NOWIRE and carried out while he visited ETH Zurich.

¹Noncoherent capacity denotes capacity in the setting where transmitter and receiver know the channel statistics but neither of them is aware of the channel realizations.

i th row and j th column is denoted by $a_{i,j}$ and we write \mathbf{a}_i^T for its i th row. If $\mathcal{I} \subseteq [1:M]$, we denote by $\mathbf{A}_{\mathcal{I}} \in \mathbb{C}^{|\mathcal{I}| \times N}$ the submatrix of \mathbf{A} obtained by retaining all rows \mathbf{a}_i^T of \mathbf{A} with row index $i \in \mathcal{I}$. Similarly, for a vector $\mathbf{x} \in \mathbb{C}^M$ we denote its i th entry by x_i . If $\mathcal{I} \subseteq [1:M]$ we denote by $\mathbf{x}_{\mathcal{I}} \in \mathbb{C}^{|\mathcal{I}|}$ the vector obtained by retaining the entries x_i of \mathbf{x} with $i \in \mathcal{I}$. We write \mathbf{e}_i for the i th unit vector of appropriate size and \mathbf{I}_M for the identity matrix of size $M \times M$. For a vector \mathbf{x} , $\text{diag}(\mathbf{x})$ denotes the diagonal matrix that has the entries of \mathbf{x} in its main diagonal. For two matrices \mathbf{A} and \mathbf{B} of arbitrary size, $\text{diag}(\mathbf{A}, \mathbf{B})$ is the 2×2 block matrix that has the matrix \mathbf{A} as upper left block, \mathbf{B} as lower right block, and $\mathbf{0}$ as upper right and lower left block. For N matrices $\mathbf{A}_1, \dots, \mathbf{A}_N$, we define $\text{diag}(\mathbf{A}_1, \dots, \mathbf{A}_N) \triangleq \text{diag}(\text{diag}(\mathbf{A}_1, \dots, \mathbf{A}_{N-1}), \mathbf{A}_N)$. We designate the Kronecker product of the matrices \mathbf{A} and \mathbf{B} as $\mathbf{A} \otimes \mathbf{B}$; to simplify notation, we use the convention that the ordinary matrix product precedes the Kronecker product, i.e., $\mathbf{AB} \otimes \mathbf{C} \triangleq (\mathbf{AB}) \otimes \mathbf{C}$. For a function f , we write $f \neq 0$ if there exists a vector \mathbf{x} in the domain of f such that $f(\mathbf{x}) \neq 0$. For two functions f and g , the notation $f = \mathcal{O}(g)$ means that $\limsup_{x \rightarrow \infty} (|f(x)/g(x)|) < \infty$. If $x \in \mathbb{R}$, $\lfloor x \rfloor \triangleq \max\{m \in \mathbb{Z} \mid m \leq x\}$ and $\lceil x \rceil \triangleq \min\{m \in \mathbb{Z} \mid m \geq x\}$ with \mathbb{Z} denoting the set of integers. The logarithm to the base 2 is written as $\log(\cdot)$. The expectation operator is denoted by $\mathbb{E}[\cdot]$. Finally, $\mathcal{CN}(\mathbf{a}, \mathbf{C})$ stands for the distribution of a jointly proper Gaussian random vector with mean \mathbf{a} and covariance matrix \mathbf{C} .

II. SYSTEM MODEL

We consider a SIMO channel with R receive antennas. The fading in each component SISO channel follows a correlated block-fading model [2], with input-output relation for a given block

$$\mathbf{y}_m = \sqrt{\rho} \text{diag}(\mathbf{h}_m) \mathbf{x} + \mathbf{n}_m, \quad m \in [1:R] \quad (1)$$

where ρ denotes the signal-to-noise ratio (SNR), $\mathbf{x} \in \mathbb{C}^L$ is the transmitted signal vector, $\mathbf{y}_m \in \mathbb{C}^L$ is the received signal vector corresponding to the m th receive antenna, and $\mathbf{n}_m \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_L)$ is additive noise. Finally, $\mathbf{h}_m \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}\mathbf{Q}^H)$ is the vector of channel coefficients between the transmit antenna and the m th receive antenna. Here,² $\mathbf{Q} \in \mathbb{C}^{L \times Q}$ and $R \leq Q \triangleq \text{rank}(\mathbf{Q}) < L$.

Without loss of generality, we assume that the row vectors \mathbf{q}_i^T of \mathbf{Q} satisfy $\mathbf{q}_i \neq \mathbf{0}$ ($i \in [1:L]$). The vectors \mathbf{h}_m and \mathbf{n}_m are assumed to be mutually independent and independent across $m \in [1:R]$. It will turn out to be convenient to write the channel-coefficient vector in whitened form as $\mathbf{h}_m = \mathbf{Q}\mathbf{s}_m$, where $\mathbf{s}_m \in \mathbb{C}^Q$ with $\mathbf{s}_m \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_Q)$. Finally, we assume that \mathbf{s}_m and \mathbf{n}_m change in an independent fashion from block to block for all $m \in [1:R]$.

Setting $\mathbf{y}^T \triangleq (\mathbf{y}_1^T, \dots, \mathbf{y}_R^T)$, $\mathbf{s}^T \triangleq (\mathbf{s}_1^T, \dots, \mathbf{s}_R^T)$, $\mathbf{n}^T \triangleq (\mathbf{n}_1^T, \dots, \mathbf{n}_R^T)$, and $\mathbf{X} \triangleq \text{diag}(\mathbf{x})$, we can combine the

²When $Q = L$, capacity is known to grow double-logarithmically in SNR [7], and, hence, the pre-log is equal to zero. If $R > Q$ we can always achieve the same pre-log as for $R = Q$ by simply using only Q receive antennas.

individual input-output relations in (1) into the overall input-output relation

$$\mathbf{y} = \sqrt{\rho} \bar{\mathbf{y}} + \mathbf{n}, \quad \text{with } \bar{\mathbf{y}} \triangleq (\mathbf{I}_R \otimes \mathbf{X}\mathbf{Q})\mathbf{s}. \quad (2)$$

III. LOWER BOUND ON THE PRE-LOG

The capacity of the channel (2) is defined as

$$C(\rho) = (1/L) \sup_{f(\mathbf{x})} I(\mathbf{x}; \mathbf{y}) \quad (3)$$

where $I(\mathbf{x}; \mathbf{y})$ denotes mutual information [8, p. 251] and the supremum is taken over all input distributions $f(\mathbf{x})$ on \mathbb{C}^L that satisfy the average power constraint $\mathbb{E}[\|\mathbf{x}\|^2] \leq L$. The pre-log is defined as $\lim_{\rho \rightarrow \infty} (C(\rho)/\log(\rho))$.

The main result of this paper is the following theorem.

Theorem 1: Suppose that \mathbf{Q} satisfies the following

Property (A): There exists a subset of indices $\mathcal{K} \subseteq [1:L]$ with cardinality

$$|\mathcal{K}| \triangleq \min(\lceil (QR - 1)/(R - 1) \rceil, L) \quad (4)$$

such that every Q row vectors of $\mathbf{Q}_{\mathcal{K}}$ are linearly independent. Then, the capacity of the SIMO channel (2) can be lower-bounded as

$$C(\rho) \geq \begin{cases} (1 - 1/L) \log(\rho) + \mathcal{O}(1), & \text{if } \frac{QR-1}{R-1} \leq L \\ R(1 - Q/L) \log(\rho) + \mathcal{O}(1), & \text{else.} \end{cases} \quad (5)$$

Remark 1: The SISO pre-log is $1 - Q/L$ [2].

Remark 2: For $Q = R$ Theorem 1 reduces to [1, Th. 1].

Remark 3: Without loss of generality, we will henceforth assume that the set \mathcal{K} in Theorem 1 is given by $\mathcal{K} = [1:|\mathcal{K}|]$. This can always be achieved by reordering the scalar input-output relations in (1).

Remark 4: The pre-log is $(1 - 1/L)$, provided that $L \geq 2Q - 1$, even if $R = 2$ only.

Remark 5: Property (A) in Theorem 1 is not restrictive and is satisfied for a broad class of correlation matrices.

Proof: Since we are interested in a capacity lower bound, we can evaluate the mutual information in (3) for an appropriate input distribution. Specifically, we take the input distribution to have entries x_i ($i \in [1:L]$) that are independent and identically distributed (i.i.d.), zero mean, unit variance, and satisfy $h(x_i) > -\infty$. This implies that [7, Lem. 6.7]

$$\mathbb{E}[\log(|x_i|)] > -\infty, \quad i \in [1:L]. \quad (6)$$

For example, we can take $x_i \sim \mathcal{CN}(0, 1)$. The mutual information $I(\mathbf{x}; \mathbf{y}) = h(\mathbf{y}) - h(\mathbf{y} | \mathbf{x})$ in (3), evaluated for any input distribution satisfying these constraints, is then lower-bounded as follows. We first upper-bound $h(\mathbf{y} | \mathbf{x})$ according to [1, Eq. (8)]

$$h(\mathbf{y} | \mathbf{x}) \leq QR \log(\rho) + \mathcal{O}(1) \quad (7)$$

and then lower-bound $h(\mathbf{y})$ as in [1, Eq. (12)]

$$h(\mathbf{y}) \geq (QR + L - \alpha) \log(\rho) + h(\mathbf{P}\bar{\mathbf{y}} | \mathbf{x}_{\mathcal{P}}) + c \quad (8)$$

where c is a constant that is independent of ρ , $h(\mathbf{P}\bar{\mathbf{y}}|\mathbf{x}_{\mathcal{P}})$ is independent of ρ , $\mathcal{P} \triangleq [1:\alpha]$ with $\alpha \in [1:L]$, and

$$\mathbf{P} \triangleq \text{diag}((\mathbf{I}_L)_{\mathcal{I}_1}, \dots, (\mathbf{I}_L)_{\mathcal{I}_R}) \quad (9)$$

for sets $\mathcal{I}_1, \dots, \mathcal{I}_R \subseteq [1:L]$ satisfying

$$\sum_{i \in [1:R]} |\mathcal{I}_i| = QR + L - \alpha. \quad (10)$$

The set \mathcal{P} can be interpreted as a set of pilot positions [1]. Combining (7) and (8), the capacity lower bound in (5) is established by choosing

$$\alpha = \begin{cases} 1, & \text{if } (QR - 1)/(R - 1) \leq L \\ QR - (R - 1)L, & \text{else,} \end{cases}$$

provided that we can find sets $\mathcal{I}_1, \dots, \mathcal{I}_R$ such that $h(\mathbf{P}\bar{\mathbf{y}}|\mathbf{x}_{\mathcal{P}}) > -\infty$. The remainder of the paper is devoted to identifying such a choice for $\mathcal{I}_1, \dots, \mathcal{I}_R$ and proving that the corresponding differential entropy $h(\mathbf{P}\bar{\mathbf{y}}|\mathbf{x}_{\mathcal{P}})$ is, indeed, finite. The main idea is to choose the sets $\mathcal{I}_1, \dots, \mathcal{I}_R$ such that $h(\mathbf{P}\bar{\mathbf{y}}|\mathbf{x}_{\mathcal{P}})$ can be related to $h(\mathbf{s}, \mathbf{x}_{\mathcal{D}}) = h(\mathbf{s}) + h(\mathbf{x}_{\mathcal{D}})$ with $\mathcal{D} \triangleq [\alpha + 1:L]$ through a deterministic one-to-one mapping. The quantity $h(\mathbf{s}) + h(\mathbf{x}_{\mathcal{D}})$ is much easier to deal with than $h(\mathbf{P}\bar{\mathbf{y}}|\mathbf{x}_{\mathcal{P}})$.

Condition (10) implies that the mapping

$$(\mathbf{s}, \mathbf{x}_{\mathcal{D}}) \mapsto \mathbf{P}\bar{\mathbf{y}} = \mathbf{P}(\mathbf{I}_R \otimes \mathbf{X}\mathbf{Q})\mathbf{s} \quad (11)$$

is between two vector spaces of the *same* dimension $QR + L - \alpha$, which is a necessary condition for this mapping to be one-to-one. Note that the RHS of (11) also depends on $\mathbf{x}_{\mathcal{P}}$, which is, however, taken to be fixed, reflecting the fact that the pilot symbols are known to both transmitter and receiver. Any dependence on \mathbf{x} will henceforth implicitly mean a dependence on $\mathbf{x}_{\mathcal{D}}$ only. We set $\mathcal{I}_R \triangleq [1:L]$ and shall choose $\mathcal{I}_1, \dots, \mathcal{I}_{R-1} \subseteq [1:L]$ as follows:

- (a) If $(QR - 1)/(R - 1) > L$, we set $\mathcal{I}_1 \triangleq \dots \triangleq \mathcal{I}_{R-1} \triangleq [1:L]$.
- (b) If $(QR - 1)/(R - 1) \leq L$, we let

$$\begin{aligned} \mathcal{I}_m &\triangleq [1:Q + k + 1], & m \in [1:R - l - 1] \\ \mathcal{I}_m &\triangleq [1:Q + k + 2], & m \in [R - l:R - 1] \end{aligned} \quad (12)$$

with $k \triangleq \lfloor (Q - R)/(R - 1) \rfloor$ and $l \triangleq Q - R - k(R - 1)$.

Now let

$$\mathbf{J}(\mathbf{s}, \mathbf{x}_{\mathcal{D}}) \triangleq \frac{\partial \mathbf{P}\bar{\mathbf{y}}}{\partial (\mathbf{s}, \mathbf{x}_{\mathcal{D}})} \quad (13)$$

be the Jacobian of the mapping in (11). If this mapping is one-to-one on $\mathbb{C}^{QR+L-\alpha}$ almost everywhere (a.e.), we can apply the change-of-variables theorem for integrals [9, Th. 7.26] in combination with [10, Th. 7.2] and find that

$$h(\mathbf{P}\bar{\mathbf{y}}|\mathbf{x}_{\mathcal{P}}) = h(\mathbf{s}, \mathbf{x}_{\mathcal{D}}) + 2\mathbb{E}[\log(|\det(\mathbf{J}(\mathbf{s}, \mathbf{x}_{\mathcal{D}}))|)].$$

The proof is then concluded by establishing that the mapping in (11) is one-to-one a.e. and

$$\mathbb{E}[\log(|\det(\mathbf{J}(\mathbf{s}, \mathbf{x}_{\mathcal{D}}))|)] > -\infty. \quad (14)$$

This requires an in-depth analysis of the Jacobian in (13), which will be carried out in the next section. ■

IV. PROPERTIES OF THE JACOBIAN

The following lemma provides important insights into the structure of the determinant of the Jacobian in (13).

Lemma 1: The Jacobian in (13) can be decomposed as

$$\mathbf{J}(\mathbf{s}, \mathbf{x}_{\mathcal{D}}) = \mathbf{J}_1(\mathbf{x}_{\mathcal{D}})\mathbf{J}_2(\mathbf{s})\mathbf{J}_3(\mathbf{x}_{\mathcal{D}}) \quad (15)$$

where

$$\mathbf{J}_1(\mathbf{x}_{\mathcal{D}}) \triangleq \mathbf{P}(\mathbf{I}_R \otimes \mathbf{X})\mathbf{P}^T \quad (16)$$

$$\mathbf{J}_2(\mathbf{s}) \triangleq \mathbf{P}[\mathbf{I}_R \otimes \mathbf{Q} \mid \mathbf{a}_{\alpha+1} \mid \dots \mid \mathbf{a}_L] \quad (17)$$

$$\mathbf{J}_3(\mathbf{x}_{\mathcal{D}}) \triangleq \text{diag}(\mathbf{I}_{QR}, \text{diag}(\mathbf{x}_{\mathcal{D}})^{-1}) \quad (18)$$

with

$$\mathbf{a}_i \triangleq (\mathbf{I}_R \otimes \text{diag}(\mathbf{e}_i)\mathbf{Q})\mathbf{s}, \quad i \in [1:L]. \quad (19)$$

Proof: The lemma follows by noting that

$$\begin{aligned} \frac{\partial \bar{\mathbf{y}}}{\partial \mathbf{x}_i} &= \frac{\partial}{\partial \mathbf{x}_i} \left(\sum_{j \in [1:L]} x_j (\mathbf{I}_R \otimes \text{diag}(\mathbf{e}_j)\mathbf{Q})\mathbf{s} \right) \\ &= \mathbf{a}_i, \quad i \in [1:L] \\ \frac{\partial \bar{\mathbf{y}}}{\partial \mathbf{s}} &= \mathbf{I}_R \otimes \mathbf{X}\mathbf{Q}. \end{aligned}$$

Based on (6), we can conclude that³

$$\begin{aligned} &\mathbb{E}[\log(|\det(\mathbf{J}_1(\mathbf{x}_{\mathcal{D}}))\det(\mathbf{J}_3(\mathbf{x}_{\mathcal{D}}))|)] \\ &= \sum_{j \in \mathcal{P}} R \log(|x_j|) + \sum_{i \in [1:R-1]} \sum_{j \in \mathcal{I}_i \setminus \mathcal{P}} \mathbb{E}[\log(|x_j|)] > -\infty. \end{aligned}$$

To conclude the proof of (14) it therefore remains to show that $\mathbb{E}[\log(|\det(\mathbf{J}_2(\mathbf{s}))|)] > -\infty$. Direct computation reveals that each vector $\mathbf{P}\mathbf{a}_i$ in (17) with $i \in \mathcal{I}_R \setminus \mathcal{I}_{R-1}$ contains only one nonzero element, which is given by $\mathbf{q}_i^T \mathbf{s}_R$. Applying the Laplace formula [11, p.7] to $\mathbf{J}_2(\mathbf{s})$ in (17) therefore yields the decomposition

$$|\det(\mathbf{J}_2(\mathbf{s}))| = |\det(\mathbf{J}_4(\mathbf{s}))| \prod_{i \in \mathcal{I}_R \setminus \mathcal{I}_{R-1}} |\mathbf{q}_i^T \mathbf{s}_R| \quad (20)$$

with

$$\mathbf{J}_4(\mathbf{s}) \triangleq \mathbf{P}_1[(\mathbf{I}_R \otimes \mathbf{Q}) \mid \mathbf{a}_{\alpha+1} \mid \dots \mid \mathbf{a}_{|\mathcal{I}_{R-1}|}] \quad (21)$$

and

$$\mathbf{P}_1 \triangleq \text{diag}((\mathbf{I}_L)_{\mathcal{I}_1}, \dots, (\mathbf{I}_L)_{\mathcal{I}_{R-1}}, (\mathbf{I}_L)_{\mathcal{I}_{R-1}}).$$

The expectation of the logarithm of the second term on the RHS in (20) is finite because $\mathbf{s}_R \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_Q)$. It remains to show that $\mathbb{E}[\log(|\det(\mathbf{J}_4(\mathbf{s}))|)] > -\infty$. This is the most technical part in the proof of Theorem 1 and can be accomplished by applying methods from algebraic geometry, namely Theorem 2 in Appendix A, which is a consequence of Hironaka's celebrated theorem on resolution of singularities [4], [5]. A direct proof would require showing that the expected log of the determinant of the (high-dimensional) matrix $\mathbf{J}_4(\mathbf{s})$ is finite, which seems exceedingly difficult. Hironaka's theorem

³We assume that $x_j \neq 0$ for all $j \in \mathcal{P}$.

drastically simplifies the proof as it tells us that $\det(\mathbf{J}_4(\cdot)) \not\equiv 0$ implies that $\mathbb{E}[\log(|\det(\mathbf{J}_4(\mathbf{s}))|)] > -\infty$. We start by noting that $\det(\mathbf{J}_4(\mathbf{s}))$ is a homogeneous polynomial in s_1, \dots, s_{QR} of degree $D \triangleq |\mathcal{I}_{R-1}| - \alpha$, i.e.,

$$\det(\mathbf{J}_4(\lambda \mathbf{s})) = \lambda^D \det(\mathbf{J}_4(\mathbf{s})), \quad \forall \lambda \in \mathbb{C}, \quad (22)$$

which allows us to apply the following proposition (for $M = QR$ and D defined above):

Proposition 1: Let g be a homogeneous polynomial in s_1, \dots, s_M of degree $D \in \mathbb{N}_0$ with $\mathbf{s} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$. Then $g \not\equiv 0$ implies that

$$\mathbb{E}[\log(|g(\mathbf{s})|)] > -\infty. \quad (23)$$

Proof: Writing \mathbf{s} as $\mathbf{s} = \|\mathbf{s}\|(\mathbf{s}/\|\mathbf{s}\|)$ and using the fact that g is a homogeneous polynomial of degree D , we can upper-bound the absolute value of the expectation in (23) by

$$|\mathbb{E}[\log(|g(\mathbf{s})|)]| \leq D \underbrace{\mathbb{E}[\log(\|\mathbf{s}\|)]}_{\triangleq A} + \underbrace{\mathbb{E}[\log(|g(\mathbf{s}/\|\mathbf{s}\|)|)]}_{\triangleq B}.$$

Then [7, Lem. 6.7] together with $\mathbf{s} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ implies that $A < \infty$. Introducing polar coordinates [12, p. 55] $r \in \mathbb{R}_+$ and

$$\boldsymbol{\theta} \triangleq (\theta_1, \dots, \theta_{2M-1})^T \in \Delta \triangleq [0, \pi]^{2M-2} \times [0, 2\pi]$$

for the complex vector $\mathbf{s} \in \mathbb{C}^M$, we can further upper-bound B according to

$$B \leq \int_0^\infty \exp(-r^2) r^{2M-1} dr \times \int_\Delta |\log(|f(\boldsymbol{\theta})|)| d\boldsymbol{\theta}$$

where f is obtained from g by changing to polar coordinates $r \in \mathbb{R}_+$, $\boldsymbol{\theta} \in \Delta$. Note that f is a real analytic function [6, Def. 1.6.1], $g \not\equiv 0$ implies that $f \not\equiv 0$, and we are integrating $|\log(|f(\cdot)|)|$ over a compact set Δ . We can therefore apply Theorem 2 in Appendix A to conclude that $B < \infty$. ■

It remains to show that $\det(\mathbf{J}_4(\cdot)) \not\equiv 0$ for our specific choice of sets $\mathcal{I}_1, \dots, \mathcal{I}_{R-1}$, which will be proved in the following lemma:

Lemma 2: Property (A) in Theorem 1 implies that $\det(\mathbf{J}_4(\cdot)) \not\equiv 0$.

Proof: See Appendix B ■

In summary, we proved that (14) holds provided that Property (A) in Theorem 1 is satisfied. It turns out that $\det(\mathbf{J}_4(\cdot)) \not\equiv 0$ also implies that the mapping in (11) is one-to-one a.e. on $\mathbb{C}^{QR+L-\alpha}$. The proof, which is along the lines of the proof of [1, Lem. 2], is omitted due to space limitations. This completes the proof of Theorem 1.

APPENDIX

A. Resolution of singularities

In this appendix, we show how Hironaka's theorem on resolution of singularities can be used to prove that $\int_{\Delta \subset \mathbb{R}^M} |\log(|f(\mathbf{x})|)| d\mathbf{x} < \infty$ provided that $f \not\equiv 0$ is a real analytic function and Δ is a compact set.

We start by defining notation that will be used in this appendix. Let $C^M(\mathbf{x}, \epsilon) \triangleq \{\mathbf{y} \in \mathbb{R}^M \mid |x_i - y_i| < \epsilon, \forall i \in [1:M]\}$ denote the open cube with side length 2ϵ centered at

\mathbf{x} . For $\mathbf{x} \in \mathbb{C}^M$ and $\mathbf{m} \in \mathbb{N}_0^M$ let $\mathbf{x}^{\mathbf{m}} \triangleq x_1^{m_1} \dots x_M^{m_M}$. If Σ is a subset of the image of a map f then $f^{-1}(\Sigma)$ denotes the inverse image of Σ .

The following lemma is an immediate consequence of a modified version of Hironaka's theorem [3, Th. 2.3]. This modified version originally appeared in [13]. The main point of this lemma is that it allows us to rewrite any real analytic function [6, Def. 1.6.1] $f \not\equiv 0$ locally as a product of a monomial and a nonvanishing real analytic function.

Lemma 3: Let f be a real analytic function from a neighborhood $\Omega \subseteq \mathbb{R}^D$ of $\mathbf{x} \in \mathbb{R}^D$ to \mathbb{R} . Suppose that $f(\mathbf{x}) = 0$. Then, there exists a triple (Σ, Ξ, ϕ) , where

- (a) $\Sigma \subseteq \Omega$ is an open set in \mathbb{R}^D with $\mathbf{x} \in \Sigma$,
- (b) Ξ is a D -dimensional real analytic manifold [3, Def. 2.10],
- (c) $\phi : \Xi \rightarrow \Sigma$ is a real analytic map

that satisfies the following conditions:

- (1) The map ϕ is proper, i.e., the inverse image of any compact set is compact.
- (2) The map ϕ is a real analytic isomorphism between $\Xi \setminus (f \circ \phi)^{-1}(0)$ and $\Sigma \setminus f^{-1}(0)$.
- (3) For each point $P \in \Xi$, there exists a coordinate chart $\{\Xi_P, \varphi_P\}$ such that $P \in \Xi_P \subseteq \Xi$, $\varphi_P : C^D(\mathbf{0}, \epsilon_P) \rightarrow \Xi_P$ is a real analytic isomorphism for some $\epsilon_P > 0$ with $\varphi_P(\mathbf{0}) = P$,

$$|f(\phi \circ \varphi_P(\mathbf{u}))| = h_P(\mathbf{u}) \mathbf{u}^{\mathbf{m}_P}, \quad \forall \mathbf{u} \in C^D(\mathbf{0}, \epsilon_P)$$

where h_P is a nonvanishing real analytic function on $C^D(\mathbf{0}, \epsilon_P)$ and $\mathbf{m}_P \in \mathbb{N}_0^D$, and the determinant of the Jacobian of the mapping $\phi \circ \varphi_P$ satisfies

$$\det\left(\frac{\partial(\phi \circ \varphi_P(\mathbf{u}))}{\partial \mathbf{u}}\right) = g_P(\mathbf{u}) \mathbf{u}^{\mathbf{n}_P}, \quad \forall \mathbf{u} \in C^D(\mathbf{0}, \epsilon_P)$$

where g_P is a nonvanishing real analytic function on $C^D(\mathbf{0}, \epsilon_P)$ and $\mathbf{n}_P \in \mathbb{N}_0^D$.

Proof: The main idea is to apply [3, Th. 2.3] to the function $\tilde{f}(\mathbf{z}) \triangleq f(\mathbf{z} + \mathbf{x})$, $\forall \mathbf{z} \in \Omega - \mathbf{x}$. We omit the details due to space limitations. ■

We are now in a position to state the theorem that is needed to prove $B < \infty$ in the proof of Proposition 1.

Theorem 2: Let $f \not\equiv 0$ be a real analytic function on an open set $\Omega \subset \mathbb{R}^D$. Then

$$\int_\Delta |\log(|f(\mathbf{x})|)| d\mathbf{x} < \infty \quad (24)$$

for all compact sets $\Delta \subset \Omega$.

Proof: Let $\mathbf{x} \in \Delta$. If $f(\mathbf{x}) = 0$ then Lemma 3 implies that there exists a triple $(\Sigma_{\mathbf{x}}, \Xi_{\mathbf{x}}, \phi_{\mathbf{x}})$ where $\Sigma_{\mathbf{x}} \subseteq \Omega$ is an open set containing \mathbf{x} , $\Xi_{\mathbf{x}}$ is a real analytic manifold, and $\phi_{\mathbf{x}} : \Xi_{\mathbf{x}} \rightarrow \Sigma_{\mathbf{x}}$ is a proper real analytic map. Moreover, for each $P \in \Xi_{\mathbf{x}}$ there exists a coordinate chart $\{\Xi_{\mathbf{x},P}, \varphi_{\mathbf{x},P}\}$ such that $\Xi_{\mathbf{x},P} = \varphi_{\mathbf{x},P}(C^D(\mathbf{0}, \epsilon_{\mathbf{x},P}))$, $\varphi_{\mathbf{x},P}(\mathbf{0}) = P$, and

$$|f \circ \phi_{\mathbf{x}} \circ \varphi_{\mathbf{x},P}(\mathbf{u})| = h_{\mathbf{x},P}(\mathbf{u}) \mathbf{u}^{\mathbf{m}_{\mathbf{x},P}} \quad (25)$$

$$\det\left(\frac{\partial \phi_{\mathbf{x}} \circ \varphi_{\mathbf{x},P}(\mathbf{u})}{\partial \mathbf{u}}\right) = g_{\mathbf{x},P}(\mathbf{u}) \mathbf{u}^{\mathbf{n}_{\mathbf{x},P}}$$

for all $\mathbf{u} \in C^D(\mathbf{0}, \epsilon_{x,P})$, where $g_{x,P}$ and $h_{x,P}$ are nonvanishing real analytic functions on $C^D(\mathbf{0}, \epsilon_{x,P})$. We can choose $\epsilon_{x,P}$ sufficiently small so that $g_{x,P}$ and $h_{x,P}$ are bounded on $C^D(\mathbf{0}, \epsilon_{x,P})$. If $f(\mathbf{x}) \neq 0$ the existence of a triple $(\Sigma_x, \Xi_x, \phi_x)$ with the properties specified above is guaranteed by taking $\Xi_x = \Sigma_x$ sufficiently small such that f does not vanish on Σ_x and by setting ϕ_x to be the identity map.

Now for each $\mathbf{x} \in \Delta$, we choose an open neighborhood Σ'_x and a compact neighborhood Δ_x such that $\mathbf{x} \in \Sigma'_x \subset \Delta_x \subset \Sigma_x$. Since Δ is a compact set, there exists a finite set of vectors $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ in Δ such that

$$\Delta \subset \bigcup_{i \in [1:N]} \Sigma'_{\mathbf{x}_i} \subset \bigcup_{i \in [1:N]} \Delta_{\mathbf{x}_i}.$$

For each $i \in [1:N]$, set $\Delta_i \triangleq \Delta_{\mathbf{x}_i}$, $\Sigma_i \triangleq \Sigma_{\mathbf{x}_i}$, $\Xi_i \triangleq \Xi_{\mathbf{x}_i}$, and $\phi_i \triangleq \phi_{\mathbf{x}_i}$. Since the mapping $\phi_i: \Xi_i \rightarrow \Sigma_i$ is proper, each set $\phi_i^{-1}(\Delta_i) \subset \Xi_i$ is a compact set. Therefore, there exists a finite set of points $\{P_1, \dots, P_{M_i}\}$ in Ξ_i such that

$$\phi_i^{-1}(\Delta_i) \subset \bigcup_{j \in [1:M_i]} \Xi_{i,j} \quad (26)$$

with $\Xi_{i,j} \triangleq \Xi_{\mathbf{x}_i, P_j}$. Since (26) holds for all $i \in [1:N]$, we can upper-bound the integral in (24) as follows:

$$\begin{aligned} \int_{\Delta} |\log(|f(\mathbf{x})|)| d\mathbf{x} &\leq \sum_{i \in [1:N]} \int_{\Delta_i} |\log(|f(\mathbf{x})|)| d\mathbf{x} \\ &\leq c_1 \sum_{i \in [1:N]} \sum_{j \in [1:M_i]} \int_{C^D(\mathbf{0}, \epsilon_{i,j})} |\mathbf{u}^{\mathbf{n}_{i,j}} \log(|\mathbf{u}^{\mathbf{m}_{i,j}}|)| d\mathbf{u} + c_2 \\ &< \infty \end{aligned}$$

where $c_1, c_2 > 0$ are positive real numbers, $g_{i,j}, h_{i,j}$ are bounded nonvanishing real analytic functions on $C^D(\mathbf{0}, \epsilon_{i,j})$, $\mathbf{m}_{i,j}, \mathbf{n}_{i,j}$ are vectors of nonnegative integers, and we changed variables according to (25). ■

B. Proof of Lemma 2

We present a proof for $\alpha = 1, l \neq 0$ and skip the (simpler) cases $\alpha > 1$ and $\alpha = 1, l = 0$.

Suppose that $\alpha = 1$ and $l \neq 0$. We can write $\mathbf{J}_4(\mathbf{s})$ in (17) as $\mathbf{J}_4(\mathbf{s}) = [\text{diag}(\mathbf{Q}_{\mathcal{I}_1}, \dots, \mathbf{Q}_{\mathcal{I}_{R-1}}, \mathbf{Q}_{\mathcal{I}_{R-1}}) \mathbf{A}]$ with $\mathbf{A}^T \triangleq [\mathbf{A}_1^T \dots \mathbf{A}_R^T]$ and \mathbf{A}_m defined as

$$\mathbf{A}_m \triangleq \begin{pmatrix} 0 & 0 & 0 & 0 \\ \mathbf{q}_2^T \mathbf{s}_m & 0 & 0 & 0 \\ \vdots & \ddots & 0 & 0 \\ 0 & \dots & \mathbf{q}_{Q+k+1}^T \mathbf{s}_m & 0 \end{pmatrix}, \quad m \in [1:R-l-1]$$

$$\mathbf{A}_m \triangleq \begin{pmatrix} 0 & 0 & 0 \\ \mathbf{q}_2^T \mathbf{s}_m & 0 & 0 \\ \vdots & \ddots & 0 \\ 0 & \dots & \mathbf{q}_{Q+k+2}^T \mathbf{s}_m \end{pmatrix}, \quad m \in [R-l:R].$$

Property (A) in Theorem 1 implies that for arbitrary subsets

$$\begin{aligned} \mathcal{K}_m &\subseteq [2:Q+k+1], \quad m \in [1:R-l-1] \\ \mathcal{K}_m &\subseteq [2:Q+k+2], \quad m \in [R-l:R] \end{aligned}$$

with $|\mathcal{K}_m| = Q-1$, we can find vectors $\mathbf{s}_m \in \mathbb{C}^Q$ ($m \in [1:R]$) such that

- (a) $\mathbf{q}_j^T \mathbf{s}_m = 0$ for all vectors \mathbf{q}_j^T with $j \in \mathcal{K}_m$;
- (b) $\mathbf{q}_j^T \mathbf{s}_m \neq 0$ for all vectors with $j \in \mathcal{K} \setminus \mathcal{K}_m$.

This implies that for each choice of such sets \mathcal{K}_m ($m \in [1:R]$), there exists a set of vectors $\mathbf{s}_m \in \mathbb{C}^Q$ ($m \in [1:R]$) such that the number of nonzero elements n_m in each matrix \mathbf{A}_m satisfies

$$\begin{aligned} n_m &= k+1, \quad m \in [1:R-l-1] \\ n_m &= k+2, \quad m \in [R-l:R]. \end{aligned} \quad (27)$$

Moreover, we have

$$\sum_{m \in [1:R]} n_m = Q + k + 1$$

which implies that we can choose the subsets \mathcal{K}_m and the vectors \mathbf{s}_m ($m \in [1:R]$) such that each column of \mathbf{A} contains precisely one nonzero element. Applying the Laplace formula [11, p. 7] iteratively, we therefore get

$$|\det(\mathbf{J}_4(\mathbf{s}))| = c \prod_{m \in [1:R]} |\det(\mathbf{Q}_{\mathcal{K}_m \cup \{1\}})| > 0$$

where c is a positive constant and we used Property (A) in Theorem 1 in the last step. ■

REFERENCES

- [1] V. I. Morgenshtern, G. Durisi, and H. Bölcskei, "The SIMO pre-log can be larger than the SISO pre-log," in *Proc. IEEE Int. Symp. Inf. Th. (ISIT 2010)*, Austin, TX, June 2010, pp. 320–324.
- [2] Y. Liang and V. V. Veeravalli, "Capacity of noncoherent time-selective Rayleigh-fading channels," *IEEE Trans. Inf. Th.*, vol. 50, no. 12, pp. 3095–3110, Dec. 2004.
- [3] S. Watanabe, *Algebraic Geometry and Statistical Learning Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2009, vol. 25.
- [4] H. Hironaka, "Resolution of singularities of an algebraic variety over a field of characteristic zero: I," *Math. Ann.*, vol. 79, no. 1, pp. 109–203, Jan. 1964.
- [5] —, "Resolution of singularities of an algebraic variety over a field of characteristic zero: II," *Math. Ann.*, vol. 79, no. 2, pp. 205–326, Mar. 1964.
- [6] S. G. Krantz and H. R. Parks, *A Primer of Real Analytic Functions*. Basel, Switzerland: Birkhäuser, 1992, vol. 4.
- [7] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels," *IEEE Trans. Inf. Th.*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003.
- [8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY: Wiley, 2006.
- [9] W. Rudin, *Real and Complex Analysis*, 3rd ed. New York, NY: McGraw-Hill, 1987.
- [10] K. Fritzsche and H. Grauert, *From Holomorphic Functions to Complex Manifolds*, 1st ed. New York, NY: Springer, 2002.
- [11] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
- [12] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York, NY: Wiley, 1982.
- [13] M. F. Atiyah, "Resolution of singularities and division of distributions," *Comm. Pure and Appl. Math.*, vol. 13, pp. 145–150, 1970.